AD-A105 951   COLORADO UNIV AT BOULDER DEPT OF COMPUTER SCIENCE   F/G 12/1
               SUMMARY OF RESEARCH UNDER ARO CONTRACT DAAG-79-C-0023.(U)
               SEP 81   R B SCHNABEL
                                              ARO-15809.3-M   DAAG29-79-C-0023
UNCLASSIFIED   NL

| OF |
AD A
105951

END
DATE
FILMED
11-81
DTIC

LEVEL II   ARO 15809.3-M

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	2 GOVT ACCESSION NO	3 RECIPIENT'S CATALOG NUMBER
	AD-A105951	

4. TITLE (and Subtitle)	5 TYPE OF REPORT & PERIOD COVERED
Summary of Research under ARO Contract DAAG-79-C-0023	Final Report, 1 June 1979 — 31 May 1981
	6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)	8 CONTRACT OR GRANT NUMBER(s)
Robert B. Schnabel	DAAG-79-C-0023

9. PERFORMING ORGANIZATION NAME AND ADDRESS	10 PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
Dept. of Computer Science University of Colo., Boulder, Colorado	

11. CONTROLLING OFFICE NAME AND ADDRESS	12 REPORT DATE
U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709	September 11, 1981
	13 NUMBER OF PAGES
	9

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15 SECURITY CLASS. (of this report)
	Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

DTIC
ELECTE
OCT 22 1981
D

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

NA

18. SUPPLEMENTARY NOTES

The view, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Unconstrained Minimization, Secant Updates, Conic Models, Tensor Models

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This research has investigated several topics in solving unconstrained minimization, systems of nonlinear equations, and indefinite linearly constrained minimization problems. In unconstrained minimization, a new projected update was developed and tested. While marginal improvements over the BFGS were found, they probably do not justify the use of the new update. The use of conic models for unconstrained minimization problems when analytic or finite difference derivatives are available was also investigated, and a new algorithm was developed and tested. The results show reasonable improvements in many cases, and indicate that conic algorithms for

81 10 20   409496

20. ABSTRACT CONTINUED

minimization should continue to be considered. Recently we have also developed a tensor algorithm for solving systems of nonlinear equations, and the preliminary computational results show considerable improvements over the corresponding standard algorithm, especially on problems where the Jacobian at the solution is singular. This approach seems to hold excellent promise. Other research in linearly constrained minimization and deriving and analyzing least change secant updates also is reported.

Accession For		
NTIS GRA&I	☒	
DTIC TAB	☐	
Unannounced	☐	
Justification		
By		
Distribution/		
Availability Codes		
Dist	Avail and/or Special	
A		

1. Introduction

Our research under this contract has mainly been concerned with the unconstrained minimization problem

$$\min_{x \in R^n} f: R^n \to R, \qquad (1.1)$$

where f is twice continuously differentiable. At times we have also considered the closely related nonlinear equations problem

$$\text{given } F: R^n \to R^n, \text{ find } x_* \in R^n \text{ such that } F(x_*) = 0. \qquad (1.2)$$

One project has been concerned with the linearly constained minimization problem

$$\min_{x \in R^n} f: R^n \to R$$
$$\text{subject to } c_i(x) \leq 0, \, i = 1, \, \dots, \, m \qquad (1.3)$$
$$c_i: R^n \to R.$$

The main topics we have investigated during this research period are:

1. Projected updates for unconstrained minimization.

2. Deriving least change secant updates for unconstrained minimization and nonlinear equations.

3. Conic models for unconstrained minimization.

4. Tensor models for systems of nonlinear equations.

5. Algorithms for indefinite linearly constrained minimization.

A minor topic investigated was

6. Convergence of secant approximations to the correct Jacobian or Hessian value.

Our findings on each of these topics are discussed briefly in Sections 2 - 7. Our work on topics 1, 2, and 6 is essentially complete, while topics 3, 4, and 5 were started towards the end of the research period, and are the main topics we are currently investigating under ARO contract DAAG29-81-K-0181,

which started on June 15, 1981.

A list of publications and technical reports during the research period is given in Section 8. However, due to our co-authorship of a research book on unconstrained minimization and nonlinear equations with J. Dennis, which will be completed in Fall 1981, there are several papers reporting our research under this contract that have not been completed. They are listed separately in Section 8. The findings of two of these papers are partially reported in two of the M.S. theses that are listed in Section 9. Most of these research results have been reported at major scientific meetings in our field; we list these presentations in Section 10.

2. Projected updates for unconstrained minimization

A new technique for deriving the BFGS update, developed by Dennis, was extended by us to derive a new projected BFGS update. (All this research is reported in Dennis and Schnabel [4].) This update is the same as Davidon's [1] projected update on quadratic functions, but is different in general. Our research assistant, M. LaRue, conducted an extensive series of tests comparing an algorithm using the new update to the same algorithm using the standard BFGS. Over the entire More-Garbow-Hillstrom [6] test set, the improvements were small; while on the average the new update did slightly better, on a fair number of problems this wasn't the case. We still plan to write up the results of this research, but in view of the added complexity of the new update, the small improvements do not seem to warrant its use in general purpose software.

To us, the most important contribution of the computational research was the techniques and intuition we developed about incorporating information from past iterates into our algorithms. Basically, we learned that at iteration k, it did not make sense to use information from iterate k-i

unless the angle between $x_k - x_{k-i}$ and the subspace spanned by $\{x_k - x_{k-j} | j=1,\ldots,i-1\}$ was large, say $45°$. Secondly, it seemed that even though the theory permitted using information from up to n past points, an upperbound of about $\sqrt{n}$ was optimal in practice. These two observations were very important to our subsequent work on conic and tensor models.

3. Deriving least change secant updates for unconstrained minimization and nonlinear equations

We have continued our work on this topic that was started in Dennis and Schnabel [2]. A paper currently in draft form by Buckley and Schnabel contains different techniques for deriving the updates in [3], and uses these techniques to derive a new weighted least change sparse symmetric secant update. Another paper in draft form by Schnabel and Toint derives the weighted symmetric sparse projection operator, the solution to

$$\min_{B \in R^{n \times n}} \quad \| W^{-1/2}(B - A) W^{-1/2} \|_F$$

subject to B symmetric and obeying a specific sparsity pattern where $A \in R^{n \times n}$ is symmetric and $W = I +$ a matrix of rank k, $k \leq n$. This operator is shown in [3] to be a key to deriving weighted least change sparse secant updates. Unfortunately, the results of Schnabel and Toint show that for $k > 1$, computing B is computationally unattractive.

4. Conic models for unconstrained minimization

This research is based on the seminal work of Davidon [2] and the subsequent work of Sorensen [8], and is mostly reported in the M.S. thesis by Stordahl (see Section 9). It concerns using the conic model

$$m(x+d) = f + \frac{g^T d}{1+p^T d} + \frac{\frac{1}{2} d^T A d}{(1+p^T d)^2} \tag{4.1}$$

where $f \in R$, $x, d, g, p \in R^n$, $A \in R^n$, in place of the standard quadratic model in minimization algorithms. Davidon's and Sorensen's work concerned using (4.1) in an algorithm where only first derivatives were available; our research has concerned the case where analytic or finite difference second derivatives are also available. We show how to use (4.1) to interpolate $f(x)$, $\nabla f(x)$ and $\nabla^2 f(x)$ at the current iterate x_k, as well as $f(x)$ at up to n past iterates x_{k-i}. In practice, at most $\sqrt{n}$ past iterates are used at any iteration. Stordahl tested an algorithm making very simple use of this model on the test set from [6], and found average improvements of 15-30% in iterations and function evaluations over the corresponding algorithm using a quadratic model, although in some cases the new algorithm performed worse. These results have been presented at two international conferences (see Section 10) and will be reported in a forthcoming paper.

5. Tensor models for systems of nonlinear equations

This research, motivated by our work on conic models for minimization, attempts to use the model

$$M(x+d) = F + Jd + \frac{1}{2}T\overset{d}{d}$$

(5.1)

where $x, d, F \in R^n$, $J \in R^{n \times n}$, $T \in R^{n \times n \times n}$, in solving the nonlinear equations problem (1.2). Using the full term T is out of the question, since it would require n^3 storage, and the solution of a system of n quadratics in n unknowns to find the root of the model at each iteration. Our research has shown how to efficiently use a small portion of T to improve the performance of nonlinear equations algorithms, especially in problems where the Jacobian at the solution is singular, without significantly increasing the storage required or the arithmetic cost per iteration. At the kth iteration, we use (5.1) to interpolate $F(x_k)$, $F'(x_k)$, and $F(x_{k-i})$ at $m \leq \sqrt{n}$ past iterates. The first two conditions require $F = F(x_k)$ and $J = F'(x_k)$; we show that

the smallest T (in the Frobenius norm) that accomplishes the last condition is a tensor of rank m that is stored using 2m n-vectors, that is, at most $2n^{3/2}$ storage. We also show that a root of M(x+d) can then be found by solving a system of m quadratics in m unknowns, which is inexpensive compared to the one $n \times n$ linear system that is solved at each iteration of any nonlinear equations algorithm. A preliminary implementation of this algorithm is reported in the M.S. thesis by Frank (see Section 9); it contains several additional important features. So far the algorithm has performed as well or better than the corresponding algorithm using a standard linear model on virtually every test problem, and on many singular problems the improvements were quite large. This research is still in a preliminary stage and is being continued on our new ARO grant. It was reported at a recent international meeting (see Section 10), where several knowledgeable people said that they thought it had very good promise.

6. Algorithms for indefinite linearly constrained minimization

Our research assistant during the last year of this contract, J. Shultz, has been conducting his Ph.D. research on this topic. It is motivated by the algorithm we introduced in [7] for finding whether a system of linear and nonlinear inequality constraints has a feasible point. This algorithm is of considerable practical interest, for example as a Phase I procedure for G.R.G. codes for constrained minimization, but to implement it successfully requires a new algorithm for indefinite linearly constrained minimization. Our research with Shultz has focused on two topics: strategies for dropping active constraints when the reduced Hessian is indefinite, and step direction strategies when the reduced Hessian is indefinite. We believe we have satisfactorily solved both problems, and at this time, Shultz is testing the new algorithm. The results will be reported in a joint paper and in his thesis.

7. Convergence of secant approximations to the correct Jacobian or Hessian
value

At the Fall 1980 SIAM meeting in Houston, we presented a series of new
examples showing that the sequence of Jacobian or Hessian approximations
generated by secant algorithms (Broyden's method, BFGS, DFP, PSB) do not
necessarily converge to the Jacobian or Hessian matrix at the solution, even
when the iterates converge to the solution q-superlinearly. These results
show that if the Jacobian is constant in some of its rows, then Broyden's
method (or any similar rank one update method) will almost always generate
a sequence of approximations that converges to incorrect values in the other,
nonlinear rows. This behavior is replicated in practice and similar behavior
is sometimes noted when some component functions are far more nonlinear than
others. For minimization, we show that the same behavior can occur if the
Hessian is constant in some rows (and columns), but if the BFGS, DFP or any
other update from the Broyden class is used, it seems far less likely to occur.
However, the PSB updates converge to an incorrect derivative value when the
BFGS and Broyden class converge correctly, and so this isolates another reason
why unweighted symmetric secant updates are inferior to weighted symmetric
secant updates for minimization. This research is partially reported in
[5], and will be a forthcoming paper.

8. Publications and technical reports produced

1. "Determining feasibility of a set of nonlinear inequality constraints,"
to appear in Math. Prog. Study on Constrained Optimization.
2. "A new derivation of symmetric positive definite secant updates," in
Nonlinear Programming 4, O. L. Mangasarian, R. R. Meyer, S. M. Robinson,
eds., Academic Press, N. Y., 1981 (with J. E. Dennis, Jr.).
3. "Comments on evaluating algorithms and codes for mathematical programming,"
to appear in Proc' ings of the Boulder COAL conference.

4. "Unconstrained optimization in 1981," to appear in the proceedings of the
    NATO ARI on Nonlinear Optimization, M. J. D. Powell, ed., Academic Press, 1982.

5. Unconstrained Optimization and Nonlinear Equations, Prentice-Hall,
    New Jersey, 1982 (with J. E. Dennis, Jr.).

The following papers currently are in nearly-final draft form:

"Long vectors for secant updates" (with A. Buckley)

"Forcing sparsity by projecting with respect to a non-diagonally
weighted Frobenius norm" (with Ph. L. Toint).

The following are some additional papers that will be forthcoming reporting
work performed under this contract:

"Projected secant updates for unconstrained minimization"

"On the convergence of Jacobian and Hessian approximations to the correct
derivative value"

"Conic models using second derivatives for unconstrained minimization"

"Tensor models for solving systems of nonlinear equations" (with P.D. Frank)

"A modular system of algorithms for unconstrained minimization"

(with B. E. Weiss and J. E. Koontz)

9. List of scientific personnel

M. LaRue, research assistant, June-July 1979, Jan.-July 1980

J. Shultz, research assistant, Sept. 1980-May 1981 (Ph.D. expected, 1982)

Masters theses supervised in conjunction with the research for this contract:

B. E. Weiss, "A modular software package for solving unconstrained non-
linear optimization problems," 1980

K. A. Stordahl, "Unconstrained minimization using conic models and exact
second derivatives," 1980

P. D. Frank, "A second-order local model for solution of systems of nonlinear
equations," 1981

10. Presentations of research conducted under this contract at scientific
    meetings

1. "Determining feasibility of a system of nonlinear inequality constraints"
   Tenth International Math Programming Symposium, Montreal, Aug. 1979 and
   ORSA-TIMS Meeting, Milwaukee, October 1979.

2. "Consideration of scaling in unconstrained optimization algorithms"
   SIAM Meeting, Denver, Nov. 1979.

3. "A derivation of the BFGS from the Broyden update and some consequences"
   by J. Dennis, coauthor, at nonlinear Programming Symposium IV,
   Madison, July 1980.

4. "On the convergence of secant approximations to the correct derivative
   value" SIAM Meeting, Houston, Nov. 1980.

5. "Unconstrained minimization using conic models and exact second derivatives"
   Mathematical Programming Conference, Oberwohlfach, W. Germany, Jan. 1981.

6. "Unconstrained optimization in 1981" and "Nonstandard models for unconstrained
   optimization and nonlinear equations" NATO ARI on Nonlinear Optimization,
   Cambridge, England, July 1981.

Bibliography

[1] W. C. Davidon (1975), "Optimally conditioned optimization algorithms without line searches," Math. Prog. 9, 1-30.

[2] W. C. Davidon (1980), "Conic approximations and collinear scalings for optimizers," Math. Prog. 17, 268-281.

[3] J. E. Dennis and R. B. Schnabel (1979), "Least change secant updates for quasi-Newton methods," SIAM Review 21, 443-459.

[4] J. E. Dennis and R. B. Schnabel (1981), "A new derivation of symmetric positive definite secant updates," in Nonlinear Programming 4, O. L. Mangasarian, R. R. Meyer and S. M. Robinson eds., Academic Press, New York.

[5] J. E. Dennis and R. B. Schnabel (1982), "Unconstrained Optimization and Nonlinear Equations," Prentice Hall, New Jersey.

[6] J. J. Moré, B. S. Garbow and K. E. Hillstrom (1981), "Testing unconstrained optimization software," TOMS 7, 17-42.

[7] R. B. Schnabel (1981), "Determining feasibility of a set of nonlinear inequality constraints," to appear in Math. Prog. Study on Constrained Optimization.

[8] D. C. Sorenson (1980), "The q-superlinear convergence of a collinear scaling algorithm for unconstrained optimization," SIAM J. Num. Anal., 84-114.

ATE

LME